第45卷第4期 上海师范大学学报(自然科学版) Vol. 45, No. 4
2016年8月 Journal of Shanghai Normal University(Natural Sciences) Aug., 2016

DOI:10.3969/J. ISSN. 1000-5137.2016.04.001

Ridge-type spectral decomposition estimators in mixed effects models with stochastic restrictions

ZHENG Lu[1], YUE Rongxian[1], CHENG Jing[2]

(1. College of Mathematics and Science, Shanghai Normal University, Shanghai 200234, China;

2. College of Science, Anhui Agricultural University, Hefei 230036, China)

Abstract: This paper proposes a new estimation of fixed effects in linear mixed models with stochastic restrictions, which is called a conditional ridge-type spectral decomposition estimator. Using the mean squared error matrix and generalized mean squared error as criteria for comparing the estimates, we establish sufficient conditions for the superiority of the conditional ridge-type spectral decomposition estimator over the conditional spectral decomposition estimator. The upper and lower bounds of the relative efficiency are also given. Finally, a simulation example is given to illustrate the theoretical results.

Key words: linear mixed mode; mean squared error matrix; ridge-type spectral decomposition estimatior; stochastic linear restrictions

CLC number: O 212.4 **Document code**: A **Article ID**: 1000-5137(2016)04-0387-08

1 Introduction

Linear mixed model is an important statistical model. In the recent twenty years, linear mixed models have found more and more applications in the fields of biology, medicine, economy, finance, environment science, sample investigation and mechanical engineering[1-5]. Based on spectral decomposition, a further study of estimation about fixed effect is made in this paper. Spectral decomposition (SD) estimation was proposed by Wang and Yin[6]. With this method we can obtain several SD estimators, which are all unbiased estimators.

Consider the following general linear mixed model

$$y = X\beta + U\xi + \varepsilon \triangleq X\beta + u, \tag{1}$$

where y is an $n\times 1$ vector of observations, X is an $n\times p$ known design matrix with full column rank, β is an $p\times 1$ vector of fixed effect, U is an $n\times p$ design matrix, ξ is an $q\times 1$ vector of random effect, ε is an $n\times 1$ vector of random disturbances. Suppose that

$$\mathrm{E}(\xi) = 0,\quad \mathrm{E}(\varepsilon) = 0,\quad \mathrm{Cov}(\xi,\varepsilon) = 0,\quad \mathrm{Cov}(\xi) = D \geqslant 0,\quad \mathrm{Cov}(\varepsilon) = F > 0.$$

It follows that

Received date: 2016-03-20

Foundation item: Shanghai Municipal Science and Technology Research Project (14DZ1201900); NSFC grant (11471216); NSFC grant (11401056)

Corresponding author: YUE Rongxian, College of Mathematics and Science, Shanghai Normal University, No. 100 Guiling Rd, Shanghai 200234, China, E-mail: yue2@shnu.edu.cn

$$V = \mathrm{Cov}(y) = \mathrm{Cov}(u) = UDU^{\mathrm{T}} + F \geqslant 0.$$

By spectral decomposition on the covariance matrix, we have

$$V = \sum_{i=1}^{t} \lambda_i P_i,$$

where $\lambda_i (i = 1, 2, \cdots, t)$ are all the distinct eigenvalues of V, and the P_i's are the principal idempotent matrices of V corresponding to the eigenvalues λ_i with $\sum_{i=1}^{t} P_i = I$.

We multiply P_i from the left on both sides of (1), and denote

$$y^{(i)} = P_i y,\ X_i = P_i X, u_i = P_i u.$$

Then we have the following transformed model:

$$y^{(i)} = X_i \beta + u_i. \tag{2}$$

It is easy to obtain that $\mathrm{E}(u_i) = 0$ and $\mathrm{Cov}(u_i) = \lambda_i P_i$. Note that P_i is a singular matrix, and then model (2) is a singular linear model.

Wang and Yin[6] proposed spectral decomposition estimation (SDE) to estimate the fixed effects β and the variance components simultaneously. A prominent feature of this method is that for the fixed effects we can obtain several spectral decomposition estimates. Specifically, for every eigenvalues λ_i, the spectral decomposition estimators of its corresponding fixed effects β are given by

$$\hat{\beta}^i_{\mathrm{SDE}} = (X^{\mathrm{T}} P_i X)^{-} X^{\mathrm{T}} P_i y,\ i = 1, 2, \cdots, t,$$

and $c^{\mathrm{T}}\hat{\beta}^i_{\mathrm{SDE}}$ is the best linear unbiased estimator (BLUE) of fixed effects β . We can easily obtain the spectral decomposition estimator of variance component as follows

$$\hat{\lambda}_i = (y_i - X\hat{\beta}^i_{\mathrm{SDE}})^{\mathrm{T}} P_i (y_i - X\hat{\beta}^i_{\mathrm{SDE}}) / r, i = 1, 2, \cdots, t,$$

where $r = \mathrm{rk}(P_i) - \mathrm{rk}(X_i)$.

As to fixed effects, it is obvious that $\mathrm{E}(c^{\mathrm{T}}\hat{\beta}^i_{\mathrm{SDE}}) = c^{\mathrm{T}}\beta$. Note that the space spanned by the columns of X_i is contained in the space spanned by the columns of P_i, i. e. $\mathscr{M}(X_i) \subset \mathscr{M}(P_i)$. According to the unified theory of least squares by Rao[7], the BLUE for any estimable function $c^{\mathrm{T}}\beta$ is invariant with respect to the choice of generalized inverse. Thus we can choose the Moore-Penrose inverse, that is,

$$\hat{\beta}^i_{\mathrm{SDE}} = (X^{\mathrm{T}} P_i X)^{+} X^{\mathrm{T}} P_i y. \tag{3}$$

According to $\mathrm{Cov}(y) = \sum_{i=1}^{t} \lambda_i P_i$, we can easily obtain that

$$\mathrm{Cov}(\hat{\beta}^i_{\mathrm{SDE}}) = \lambda_i (X^{\mathrm{T}} P_i X)^{+}.$$

If there is an eigenvalue of $X^{\mathrm{T}} P_i X$ close to 0, the mean squared error of the SDE $\hat{\beta}^i_{\mathrm{SDE}}$ will become large abnormally. To address this issue, Yang and Li [8] defined the partial ridge-type spectral decomposition estimator as follows:

$$\hat{\beta}^i_{\mathrm{SDE}}(k_i) = (X^{\mathrm{T}} P_i X + k_i Q_i Q_i^{\mathrm{T}})^{+} X^{\mathrm{T}} P_i y, \tag{4}$$

where $k_i > 0$ and Q_i is the eigenvector matrix of $X^{\mathrm{T}} P_i X$. The partial ridge-type spectral decomposition estimator is superior to the spectral decompositionestimator in the sense of mean squared error matrix.

This paper begins with an introductory section containing a brief review of the estimators on fixed effects. It is worth noting that, the above estimators are obtained by estimating the regression coefficients β from the linear mixed model itself. Whereas in the regression model for describing economic phenomena, in addition to the sample information, we tend to get some prior information. With these prior information, the regression coefficient estimates have more superior properties than those have no prior information. The fundamental purpose of this paper is to introduce a new conditional ridge-type spectral decomposition estimator (CRSDE) for the fixed effects.

The rest of this paper is organized as follows. Section 2 introduces the conditional spectral decomposition estimator (CSDE), and Section 3 introduces the CRSDE. These two estimators are compared in Section 4. A numerical example is given to illustrate some of the theoretical results in Section 5 and some conclusion remarks are given in Section 6.

2　Conditional spectral decomposition estimate

Consider the linear mixed model (1) with respect to the following stochastic restriction:

$$r = R\beta + e, \tag{5}$$

where r is a $j\times 1$ known random vector, R is a given $j\times p$ matrix with full row rank, e is a $j\times 1$ vector of random disturbances with mean 0 and covariance matrix W which is a known positive matrix. Suppose that u and e are uncorrelated.

Merge (2) with (5) as follows:

$$\begin{pmatrix} y^{(i)} \\ r \end{pmatrix} = \begin{pmatrix} X_i^{(i)} \\ R \end{pmatrix}\beta + \begin{pmatrix} u_i \\ e \end{pmatrix}. \tag{6}$$

We denote

$$y^{(i)M} = \begin{pmatrix} y^i \\ r \end{pmatrix}, X_i^M = \begin{pmatrix} X_i \\ R \end{pmatrix}, u_i^M = \begin{pmatrix} u^i \\ e \end{pmatrix}$$

and rewrite (6) as

$$y^{(i)M} = X_i^M\beta + u_i^M. \tag{7}$$

Note that

$$\mathrm{E}(u_i^M) = 0,\quad \mathrm{Cov}(u_i^M) = \lambda_i\Sigma_i,\quad \Sigma_i = \begin{pmatrix} P_i & 0 \\ 0 & \lambda_i^{-1}W \end{pmatrix}.$$

Because (2) is a singular linear model, so the model (7) is a singular linear model with stochastic linear restrictions[9].

We define the CSDE of β to be the following:

$$\hat{\beta}^i_{\mathrm{CSDE}} = ((X_i^M)^{\mathrm{T}}\Sigma_i^- X_i^M)^+ ((X_i^M)^{\mathrm{T}}\Sigma_i^- y^{(i)M}) = (X_i^{\mathrm{T}}P_iX_i + \lambda_iR^{\mathrm{T}}W^{-1}R)^+ (X_i^{\mathrm{T}}P_iy^i + \lambda_iR^{\mathrm{T}}W^{-1}r) = (X^{\mathrm{T}}P_iX + \lambda_iR^{\mathrm{T}}W^{-1}R)^+ (X^{\mathrm{T}}P_iy + \lambda_iR^{\mathrm{T}}W^{-1}r). \tag{8}$$

Note that the estimator $\hat{\beta}^i_{\mathrm{CSDE}}$ is the BLUE of β in the linear mixed model with stochastic restrictions, and

$$\mathrm{Cov}(\hat{\beta}^i_{\mathrm{CSDE}}) = \lambda_i(X^{\mathrm{T}}P_iX + \lambda_iR^{\mathrm{T}}W^{-1}R)^+.$$

Let $\delta_{i1}\geqslant\delta_{i2}\geqslant\cdots\geqslant\delta_{ir_i}>0$ be the positive eigenvalues of $(X^{\mathrm{T}}P_iX+\lambda_iR^{\mathrm{T}}W^{-1}R)$, $r_i\leqslant p$, and $\phi_{i1},\phi_{i2},\cdots,\phi_{ir_i}$ be the corresponding standardized eigenvectors. Define $\Phi_i=(\phi_{i1},\phi_{i2},\cdots,\phi_{ir_i})$. We then have

$$X^{\mathrm{T}}P_iX + \lambda_iR^{\mathrm{T}}W^{-1}R = \Phi_i\,\mathrm{diag}(\delta_{i1},\delta_{i2},\cdots,\delta_{ir_i})\Phi_i^{\mathrm{T}} \triangleq \Phi_i\Delta_i\Phi_i^{\mathrm{T}}$$

and

$$(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R)^{+}=\Phi_i\,\mathrm{diag}(\delta_{i1}^{-1},\delta_{i2}^{-1},\cdots,\delta_{ir_i}^{-1})\Phi_i^{\mathrm T}=\Phi_i\Delta_i^{-1}\Phi_i^{\mathrm T}.$$

Therefore, we can easily get the mean squared error (MSE) of the CSDE:

$$MSE(\hat\beta^{i}_{\mathrm{CSDE}})=tr(\lambda_i(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R)^{+})=\lambda_i\sum_{j=1}^{r_i}\delta_{ij}^{-1}.$$

Obviously, when there is an eigenvalue of $(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R)$ close to zero, the MSE of the CSDE $\hat\beta^{i}_{\mathrm{CSDE}}$ will become large abnormally. Therefore, by the idea of the thought of partial ridge-type spectral decomposition estimator[8], we propose in the following section the CRSDE $\hat\beta^{i}_{\mathrm{CSDE}}(k_i)$.

3 Conditional ridge-type spectral decomposition estimate

Define the following estimators of the fixed effects β in the mixed model (7):

$$\hat\beta^{i}_{\mathrm{CSDE}}(k_i)=(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R+k_i\Phi_i\Phi_i^{\mathrm T})^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r),k_i>0,i=1,2,\cdots,t. \quad (9)$$

These are called the CRSDE. Note that

$$\hat\beta^{i}_{\mathrm{CSDE}}(k_i)=(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R+k_i\Phi_i\Phi_i^{\mathrm T})^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=$$
$$(\Phi_i\Delta_i\Phi_i^{\mathrm T}+k_i\Phi_i\Phi_i^{\mathrm T})^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=\Phi_i(\Delta_i+k_iI)^{-1}\Phi_i^{\mathrm T}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=$$
$$\Phi_i(\Delta_i+k_iI)^{-1}\Delta_i\Phi_i^{\mathrm T}\hat\beta^{i}_{\mathrm{CSDE}},$$

and then

$$|\hat\beta^{i}_{\mathrm{CSDE}}(k_i)|^2=|\Phi_i(\Delta_i+k_iI)^{-1}\Delta_i\Phi_i^{\mathrm T}\hat\beta^{i}_{\mathrm{CSDE}}|^2<|\hat\beta^{i}_{\mathrm{CSDE}}|^2.$$

We can conclude that $\hat\beta^{i}_{\mathrm{CSDE}}(k_i)$ is condensed by $\hat\beta^{i}_{\mathrm{CSDE}}$ to the origin, and

$$\mathrm E(\hat\beta^{i}_{\mathrm{CSDE}}(k_i))\neq\beta.$$

Therefore, the CRSDEs in (9) are Stein-type biased estimators.

Similar to the principal components regression estimation, we transform model (7) by $\beta=\Phi_i\gamma$, $X_i^M=Z_i\Phi_i^{\mathrm T}$. This is actually similar to reduced-order model transformation of the principal component, reducing the dimensionality from p to r_i. The model (7) is transformed to

$$y^{(i)M}=Z_i\gamma+u_i^M,\ u_i^M\sim(0,\lambda_i\Sigma_i)\ i=1,2,\cdots,t. \quad (10)$$

The CSDE of γ in (10) is given by

$$\hat\gamma^{i}_{\mathrm{CSDE}}=\Phi_i^{\mathrm T}\hat\beta^{i}_{\mathrm{CSDE}}=\Phi_i^{\mathrm T}(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R)^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=$$
$$\Phi_i^{\mathrm T}\Phi_i\Delta_i\Phi_i^{\mathrm T}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=\Delta_i^{-1}Z_i^{\mathrm T}\Sigma_i^{-}y^{(i)M}, \quad (11)$$

and the CRSDE of γ is given by

$$\hat\gamma^{i}_{\mathrm{CSDE}}(k_i)=\Phi_i^{\mathrm T}\hat\beta^{i}_{\mathrm{CSDE}}(k_i)=\Phi_i^{\mathrm T}(X^{\mathrm T}P_iX+\lambda_iR^{\mathrm T}W^{-1}R+k_i\Phi_i\Phi_i^{\mathrm T})^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=$$
$$\Phi_i^{\mathrm T}(\Phi_i\Delta_i\Phi_i^{\mathrm T}+k_i\Phi_i\Phi_i^{\mathrm T})^{+}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=(\Delta_i+k_iI)^{-1}\Phi_i^{\mathrm T}(X^{\mathrm T}P_iy+\lambda_iR^{\mathrm T}W^{-1}r)=$$
$$(\Delta_i+k_iI)^{-1}\Delta_i\hat\gamma^{i}_{\mathrm{CSDE}}\triangleq M_i\hat\gamma^{i}_{\mathrm{CSDE}}, \quad (12)$$

where $M_i=(\Delta_i+k_iI)^{-1}\Delta_i$. In the next section we discuss the estimators $\hat\gamma^{i}_{\mathrm{CSDE}}$ and $\hat\gamma^{i}_{\mathrm{CSDE}}(k_i)$.

4 Superiority of the conditional ridge-type spectral decomposition estimator

This section discusses the statistical properties of CRSDE and CSDE. The two estimators are compared under the mean squared error matrix (MSEM) and generalized mean squared error (GMSE). Here, the MSEM of an estimator $\hat{\theta}$ for a p-dimensional unknown parameter θ is defined as

$$MSEM(\hat{\theta}) = \mathrm{E}(\hat{\theta}-\theta)(\hat{\theta}-\theta)^{\mathrm{T}} = \mathrm{Cov}(\hat{\theta}) + Bias(\hat{\theta})Bias(\hat{\theta})^{\mathrm{T}},$$

where $\mathrm{Cov}(\hat{\theta})$ is the covariance matrix of $\hat{\theta}$ and $Bias(\hat{\theta}) = \mathrm{E}(\hat{\theta}) - \theta$. The GMSE of $\hat{\theta}$ is defined as

$$GMSE(\hat{\theta}) = \mathrm{E}(\hat{\theta}-\theta)^{\mathrm{T}}D(\hat{\theta}-\theta),$$

where D is a positive definite matrix. The following two lemmas are useful for comparing the MSEM and GMSE of estimators.

Lemma 4.1[10, Theorem A.71] Let A be an $n\times n$ symmetric matrix, x be an n-vector, and $\alpha>0$ be any scalar. Then the following statements are equivalent:

(i) $\alpha A - xx^{\mathrm{T}} \geqslant 0$.

(ii) $A\geqslant 0$, $x\in\mathscr{R}(A)$, and $x^{\mathrm{T}}A^{-}x\leqslant\alpha$, with A^{-} being any g-inverse of A, where $\mathscr{R}(A)$ is the vector space spanned by the column vectors of A.

Lemma 4.2[10, Theorem 3.11] Suppose $\hat{\theta}_i$, $i=1,2$ are two estimators of θ. The following two statements are equivalent:

(i) $MSEM(\hat{\theta}_1)\geqslant MSEM(\hat{\theta}_2)$.

(ii) $GMSE(\hat{\theta}_1)\geqslant GMSE(\hat{\theta}_2)$ for every positive definite matrix D.

We now calculate the MSEMs of the estimators $\hat{\gamma}^i_{\mathrm{CSDE}}$ and $\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)$ which are given in (11) and (12), respectively. The covariance matrix and bias of $\hat{\gamma}^i_{\mathrm{CSDE}}$ are the followings:

$$\mathrm{Cov}(\hat{\gamma}^i_{\mathrm{CSDE}}) = \mathrm{Cov}(\Delta_i^{-1}Z_i^{\mathrm{T}}\Sigma_i^{-}y^{(i)M}) = \Delta_i^{-1}Z_i^{\mathrm{T}}\Sigma_i^{-}\mathrm{Cov}(y^{(i)M})\Sigma_i^{-}Z_i\Delta_i^{-1} =$$

$$\lambda_i\Delta_i^{-1}Z_i^{\mathrm{T}}\Sigma_i^{-}Z_i\Delta_i^{-1} = \lambda_i\Delta_i^{-1},\ Bias(\hat{\gamma}^i_{\mathrm{CSDE}}) = \mathrm{E}(\hat{\gamma}^i_{\mathrm{CSDE}}) - \gamma = 0,$$

and then the MSEM of $\hat{\gamma}^i_{\mathrm{CSDE}}$ is

$$MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}) = \mathrm{E}(\hat{\gamma}^i_{\mathrm{CSDE}}-\gamma)(\hat{\gamma}^i_{\mathrm{CSDE}}-\gamma)^{\mathrm{T}} = \lambda_i\Delta_i^{-1}. \tag{13}$$

The covariance matrix and bias of $\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)$ are as follows:

$$\mathrm{Cov}(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \mathrm{Cov}(M_i\hat{\gamma}^i_{\mathrm{CSDE}}) = M_i\mathrm{Cov}(\hat{\gamma}^i_{\mathrm{CSDE}})M_i^{\mathrm{T}} = \lambda_iM_i\Delta_i^{-1}M_i^{\mathrm{T}},$$

$$Bias(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \mathrm{E}(M_i\hat{\gamma}^i_{\mathrm{CSDE}}) - \gamma = (M_i - I)\gamma = -k_i(\Delta_i + k_iI)^{-1}\gamma,$$

and then the MSEM of $\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)$ is

$$MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \lambda_iM_i\Delta_i^{-1}M_i^{\mathrm{T}} + (M_i - I)\gamma\gamma^{\mathrm{T}}(M_i - I)^{\mathrm{T}}. \tag{14}$$

Theorem 4.1 For model (7), if the condition $0<k_i\leqslant\dfrac{2\lambda_i}{\gamma^{\mathrm{T}}\gamma}$ holds, then we have

$$MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) \leqslant MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}).$$

Proof From expressions in (13) and (14) for $MSEM(\hat{\gamma}^i_{\mathrm{CSDE}})$ and $MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i))$, it follows that

$$MSE(\hat{\gamma}^i_{\mathrm{CSDE}}) - MSE(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \lambda_i\Delta_i^{-1} - \lambda_iM_i\Delta_i^{-1}M_i^{\mathrm{T}} - (M_i - I)\gamma\gamma^{\mathrm{T}}(M_i - I)^{\mathrm{T}} =$$

$$(\Delta_i + k_iI)^{-1}[\lambda_i(2k_iI + k_i^{2}\Delta_i^{-1}) - k_i^{2}\gamma\gamma^{\mathrm{T}}](\Delta_i + k_iI)^{-1}.$$

By Lemma 4.1, the condition $0<k_i\leqslant\dfrac{2\lambda_i}{\gamma^{\mathrm{T}}\gamma}$ implies that

$$2\lambda_i k_i I - k_i^2 \gamma\gamma^{\mathrm{T}} \geqslant 0.$$

Therefore, we have

$$\lambda_i (2k_i I + k_i^2 \Delta_i^{-1}) - k_i^2 \gamma\gamma^{\mathrm{T}} \geqslant 0$$

and the proof is completed.

Corollary 4.1 For model (7), if $0 < k_i \leqslant \frac{2\lambda_i}{\gamma^{\mathrm{T}}\gamma}$, then for any given positive matrix D we have

$$GMSE(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) \leqslant GMSE(\hat{\gamma}^i_{\mathrm{CSDE}}).$$

Theorem 4.2 For model (7) within the ellipsoid $\gamma^{\mathrm{T}} \Lambda_i \gamma \leqslant \lambda_i$ for each $i \in \{1, \cdots, t\}$, we have

$$GMSE(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) \leqslant GMSE(\hat{\gamma}^i_{\mathrm{CSDE}}), \tag{15}$$

where

$$\Lambda_i = \mathrm{diag}\left(\frac{\delta_{i1} k_i^2}{2k_i\delta_{i1} + k_i^2}, \frac{\delta_{i2} k_i^2}{2k_i\delta_{i2} + k_i^2}, \cdots, \frac{\delta_{ir_i} k_i^2}{2k_i\delta_{ir_i} + k_i^2}\right).$$

Proof From (13) and (14), we conclude that

$$MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}) - MSEM(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \lambda_i \Delta_i^{-1} - \lambda_i M_i \Delta_i^{-1} M_i^{\mathrm{T}} - (M_i - I)\gamma\gamma^{\mathrm{T}}(M_i - I)^{\mathrm{T}} =$$
$$\lambda_i (\Delta_i^{-1} - M_i \Delta_i^{-1} M_i^{\mathrm{T}}) - k_i (\Lambda_i + k_i I)^{-1} \gamma\gamma^{\mathrm{T}} (k_i (\Lambda_i + k_i I)^{-1})^{\mathrm{T}}.$$

Denote

$$A_i = \Delta_i^{-1} - M_i \Delta_i^{-1} M_i^{\mathrm{T}} = \mathrm{diag}\left(\frac{1}{\delta_{i1}} - \frac{\delta_{i1}}{(\delta_{i1} + k_i)^2}, \frac{1}{\delta_{i2}} - \frac{\delta_{i2}}{(\delta_{i2} + k_i)^2}, \cdots, \frac{1}{\delta_{ir_i}} - \frac{\delta_{ir_i}}{(\delta_{ir_i} + k_i)^2}\right),$$

$$A_i^{*} = k_i (\Delta_i + k_i I)^{-1} = \mathrm{diag}\left(\frac{k_i}{\delta_{i1} + k_i}, \frac{k_i}{\delta_{i2} + k_i}, \cdots, \frac{k_i}{\delta_{ir_i} + k_i}\right).$$

We then have

$$MSE(\hat{\gamma}^i_{\mathrm{CSDE}}) - MSE(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) = \lambda_i A_i - A_i^{*} \gamma\gamma^{\mathrm{T}} A_i^{*}.$$

Note that the condition $\gamma^{\mathrm{T}} \Lambda_i \gamma \leqslant \lambda_i$ implies that $\gamma^{\mathrm{T}} A_i^{*} A_i^{-1} A_i^{*} \gamma \leqslant \lambda_i$. According to Lemma 4.1, we have

$$MSE(\hat{\gamma}^i_{\mathrm{CSDE}}) - MSE(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)) \geqslant 0,$$

and then the desired result (15) follows from Lemma 4.2.

Now we consider the relative efficiency of the estimator $\hat{\gamma}^i_{\mathrm{CSDE}}$ with respect to the estimator $\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)$. The relative efficiency is defined by the following formula:

$$e(\hat{\gamma}^i_{\mathrm{CSDE}}) = \frac{\det(\mathrm{Cov}(\hat{\gamma}^i_{\mathrm{CSDE}}(k_i)))}{\det(\mathrm{Cov}(\hat{\gamma}^i_{\mathrm{CSDE}}))}.$$

Theorem 4.3 The relative efficiency satisfies

$$\left(\frac{\delta_{ir_i}}{\delta_{ir_i} + k_i}\right)^{r_i} \leqslant e(\hat{\gamma}^i_{\mathrm{CSDE}}) \leqslant \left(\frac{\delta_{i1}}{\delta_{i1} + k_i}\right)^{r_i}.$$

Proof According to the definition of the relative efficiency, we have

$$e(\hat{\gamma}^i_{\mathrm{CSDE}}) = \frac{\det(\lambda_i M_i \Delta_i^{-1} M_i^{\mathrm{T}})}{\det(\lambda_i \Delta_i^{-1})} = \det(M_i M_i^{\mathrm{T}}) = \det((\Delta_i + k_i I)^{-1} \Delta_i^2 (\Delta_i + k_i I)^{-1}) = \prod_{j=1}^{r_i} \left(\frac{\delta_{ij}}{\delta_{ij} + k_i}\right)^2.$$

Because $\delta_{i1} \geqslant \delta_{i2} \geqslant \cdots \geqslant \delta_{ir_i} > 0$ and

$$\frac{\delta_{ir_i}}{\delta_{ir_i} + k_i} \leqslant \frac{\delta_{ij}}{\delta_{ij} + k_i} \leqslant \frac{\delta_{i1}}{\delta_{i1} + k_i},$$

it follows that

$$\left(\frac{\delta_{ir_i}}{\delta_{ir_i}+k_i}\right)^{r_i} \leqslant e(\hat{\gamma}^i_{\mathrm{CSDE}}) \leqslant \left(\frac{\delta_{i1}}{\delta_{i1}+k_i}\right)^{r_i}.$$

The proof is completed.

5　Monte-carlo simulation study

To illustrate our theoretical results, we now consider a simulation study to compare the performance of the estimators introduced in previous sections. This study was discussed by Gumedze and Dunne[11] and Yang, Ye and Xue[12]. Here the linear mixed model is given by

$$y_{ij} = \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \xi_i + e_{ij},$$

where $i=1,\cdots,5, j=1,\cdots,10$ with $\beta=(\beta_1,\beta_2,\beta_3)^{\mathrm{T}}=(2,1.5,2)^{\mathrm{T}}$, and the ξ are iid random effects with distribution $N(0,0.5)$, and e_{ij} are iid random disturbances with distribution $N(0,1)$. The explanatory variables x_{1ij}, x_{2ij} and x_{3ij} are generated pseudo-numbers from uniform distributions $U(1,3)$, $U(2,4)$ and $U(0,1)$, respectively. The covariance matrix of $\xi_i + e_{ij}$ is given by

$$V = 0.5(I_5 \otimes e_{10}e_{10}^{\mathrm{T}}) + I_{50}.$$

The distinct eigenvalues of V are $\lambda_1=6$ and $\lambda_2=1$.

Assume that the following stochastic linear restrictions are used:

$$r = R\beta + e, R = (1,0,1), e \sim \mathrm{N}(0,0.5).$$

In the simulation study, $J=1000$ replicates are generated andthe estimated mean squared errors (EMSE) for estimators are calculated as

$$EMSE(\hat{\beta}) = \frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{3}(\hat{\beta}_{ij}-\beta_i)^2,$$

where the subscript j represents the estimators in the jth repeated experiment. Then using the equations in (3), (8) and (9) corresponding to λ_2, we compute the EMSE values of the SDE, CSDE and CRSDE by the above formulas. The simulation results are shown in Table 1.

Table 1　EMSEs of SDE, CSDE and CRSDE ($i=2$) for $\beta=(\beta_1,\beta_2,\beta_3)^{\mathrm{T}}$

$\hat{\beta}$	$k=0.05$	$k=0.15$	$k=0.25$	$k=0.35$	$k=0.45$	$k=0.55$	$k=0.65$
$\hat{\beta}^2_{\mathrm{SDE}}$	0.4165	0.4165	0.4165	0.4165	0.4165	0.4165	0.4165
$\hat{\beta}^2_{\mathrm{CSDE}}$	0.3103	0.3103	0.3103	0.3103	0.3103	0.3103	0.3103
$\hat{\beta}^2_{\mathrm{CSDE}}(k)$	0.3065	0.3015	0.2997	0.3007	0.3043	0.3103	0.3184

We observe that the CRSDE and CSDE for β are superior to the SDE. And when ridge parameter k is small enough, the CRSDE is superior to the CSDE. The Monte Carlo simulations agree with our theoretical discovery in this paper. We can conclude that the CRSDE is meaningful in practice.

6　Conclusion

In this paper, the CSDE and CRSDE for the parameters of fixed effects in a linear mixed model are proposed when the prior information is available about the parameters. Furthermore, we show that the CRSDE is superior to the CSDE and SDE in the sense of MSEM under certain conditions. The upper and lower bounds of the relative efficiency are also given. Finally, we illustrate our results with a Monte-Carlo simulation study.

References:

[1] Verbeke G, Molenberghs G. Linear mixed models in practice: a SAS-oriented approach. Lecture Notes in Statistics 126 [M]. New York: Springer-Verlag, 1997.

[2] Verbeke G, Molenberghs G. Linear mixed models for longitudinal data [M]. New York: Springer-Verlag, 2000.

[3] Wang S G, Chow S C. Advanced linear models [M]. New York: Marcel Dekker Inc, 1994.

[4] Khunri A I, Mathew T, Sinha B K. Statistical tests for mixed linear models [M]. New York: John Wiley, 1998.

[5] Searle S R, Casella G, McCulloch C E. Variance components [M]. New York: John Wiley, 1992.

[6] Wang S G, Yin S J. A new estimate of the parameters in linear mixed models [J]. Science in China (Series A), 2002, 32(5): 434-443.

[7] Rao C R, Toutenburg H. Linear models: Least squares and alternatives [M]. New York: Springer-Verlag, 1995.

[8] Yang H, Li Y L. Partial ridge-type spectral decomposition estimator in linear mixed model (in Chinese) [J]. Chinese Journal of Applied Probability, 2008, 24(3): 289-296.

[9] Xu J W, Yang H. Estimation in singular linear models with stochastic linear restrictions [J]. Communications in Statistics-Theory and Methods, 2007, 40(24): 4364-4371.

[10] Rao C R, Toutenburg H, Heumann S C. Linear models and generalizations: least squares and alternatives [M]. New York: Springer-Verlag, 2008.

[11] Gumedze F N, Dunne T T. Parameter estimation and inference in the linear mixed model [J]. Linear Algebra and its Applications, 2011, 435(8): 1920-1944.

[12] Yang H, Ye H L, Xue Kai. A further study of predictions in linear mixed models [J]. Communications in Statistics-Theory and Methods, 2014, 43(20): 4241-4252.

具有随机约束的混合效应模型参数的岭型谱分解估计

郑 鹭[1], 岳荣先[1], 程 靖[2]

(1. 上海师范大学 数理学院,上海 200234; 2. 安徽农业大学 理学院,合肥 230036)

摘 要: 对于具有随机线性约束的线性混合效应模型参数提出一种称之为条件岭型谱分解估计的方法. 利用均方误差矩阵和广义均方误差对固定效应参数的几种估计量进行比较,给出条件岭型谱分解估计优于条件谱分解估计的充分条件,并给出这两种估计的相对效率的上下界. 最后,模拟算例验证了理论结果的正确性.

关键词: 混合效应模型; 均方误差矩阵; 岭型谱分解估计; 随机线性约束

(责任编辑:冯珍珍)